CLAIMS

What is claimed is:

1. A request tracking data prefetch apparatus for a computer system, comprising:

a prefetcher coupled to a high latency memory for a processor of the computer system;

a tracker coupled to the prefetcher and configured to recognize an access to a plurality of cache lines by a processor of the computer system, and use a bit vector to predictively load a target cache line of the high latency memory into a low latency memory for the processor.

2. The apparatus of claim 1, wherein the tracker includes a tag configured to recognize an access to a corresponding plurality of cache lines of the high latency memory by the processor.

3. The apparatus of claim 2, wherein a plurality of accesses by the processor to the high latency memory as recognized by the tag is used by the tracker to determine the target cache line for a predictive load into the low latency memory.

4. The apparatus of claim 3, wherein consecutive accesses by the processor to adjacent cache lines of the high latency memory is used to determine the target cache line for a predictive load into the low latency memory.

5. The apparatus of claim 1, wherein the high latency memory comprises a memory block of a plurality of memory blocks of the computer system.

6. The apparatus of claim 5, wherein the memory block comprises a four kilobyte page of system memory of the computer system.

7. The apparatus of claim 5, wherein the tracker includes a tag configured to monitor a sub portion of the memory block for accesses by the processor.

8. The apparatus of claim 1, wherein the high latency memory is a system memory of the computer system.

9. A request tracking data prefetch apparatus for a computer system, comprising:

a processor;

a system memory coupled to the processor;

a prefetch unit coupled to the system memory;

a plurality of trackers included in the prefetch unit, wherein the trackers are respectively configured to recognize accesses to pages of the system memory; and

a cache memory coupled to the prefetch unit, wherein the prefetch unit uses a bit vector to predictively load target cache lines from the system memory into the cache memory to reduce an access latency of the processor, and wherein the target cache lines are indicated by the trackers.

10. The apparatus of claim 9, wherein each of the trackers include a tag to recognize accesses to cache lines by the processor.

11. The apparatus of claim 9, wherein a plurality of system memory accesses by the processor are used by the trackers to determine the target cache lines for a predictive load into the cache memory.

12. The apparatus of claim 11, wherein consecutive accesses by the processor to adjacent cache lines of a page is used to determine the target cache line for a predictive load into the cache memory.

13. The apparatus of claim 9, wherein the system memory comprises a plurality of 4KB pages.

14. The apparatus of claim 9, wherein each of the plurality of trackers are configured to monitor a sub portion of a page for accesses by the processor.

15. The apparatus of claim 14, wherein the cache lines are 128 byte cache lines and wherein a tag is used to monitor half of a page for accesses by the processor.

16. The apparatus of claim 9, wherein the cache memory is a prefetch cache memory within the prefetch unit.

17. The apparatus of claim 9, wherein the cache memory is an L2 cache memory.

18. A method for request tracking data prefetching for a computer system, comprising:

monitoring data transfers between a high latency memory and a low latency memory coupled to a processor by using a prefetcher;

using a bit vector to track a data transfer pattern between the high latency memory and the low latency memory;

prefetching data from the high latency memory to the low latency memory in accordance with the pattern and reducing a data access latency of a processor of the computer system.

19. The method of claim 18 wherein the computer system includes a plurality of processors, and wherein each of the processors is coupled to a respective high latency memory and a low latency memory.

20. The apparatus of claim 18, wherein consecutive accesses by the processor to adjacent cache lines of the high latency memory is used to determine a target cache line for a prefetching to the low latency memory.